CS 229r - Spectral Graph Theory in Computer Science	Prof. Salil Vadhan
Problem Set 1	
Harvard SEAS - Fall 2020	*Due: Fri. Sep. 25, 2020 (5pm)*

Your problem set solutions must be typed (in e.g. LaTeX) and submitted on Gradescope. You are allowed 12 late days for the semester, of which at most 5 can be used on any individual problem set. (1 late day = 24 hours exactly). Please name your file `ps1-lastname.*`.

The problem sets may require a lot of thought, so be sure to start them early. You are encouraged to discuss the course material and the homework problems with each other in small groups (2-3 people). Identify your collaborators on your submission. Discussion of homework problems may include brainstorming and verbally walking through possible solutions, but should not include one person telling the others how to solve the problem. In addition, each person must write up their solutions independently, and these write-ups should not be checked against each other or passed around.

Strive for clarity and conciseness in your solutions, emphasizing the main ideas over low-level details. Do not despair if you cannot solve all the problems! Difficult problems are included to stimulate your thinking and for your enjoyment, not to overwork you.

Problem 1.(Smallest possible λ_2) In class, we discussed the largest possible value of λ_2. In this problem, we will study the smallest possible nonzero value of λ_2.

1. Let G be an undirected graph whose Laplacian is L, with second-smallest eigenvalue λ_2. We know that if G is connected then $\lambda_2 > 0$. Prove that if G is connected and unweighted, then in fact $\lambda_2 \geq 1/\mathrm{poly}(n)$ by analyzing the Rayleigh quotient on all test vectors $x \perp \mathbf{1}$. For extra credit, prove the tight bound $\lambda_2 \geq 1/O(rn) \geq 1/O(n^2)$, where r is the *diameter* of the graph (i.e. the maximum shortest-path distance between pairs of vertices in the graph), and further that when G is a simple d-regular graph, we have $\lambda_2 \geq d/O(n^2)$. Note that in Lemma 10.6.1, Spielman gives a slick proof of the bound $\lambda_2 \geq 2/r \cdot (n-1)$ using tools that we have not yet covered; here we'd like you to give a "direct proof" of a similar bound by analyzing the Rayleigh quotient.

2. Prove that the bounds $1/O(rn)$ and $d/O(n^2)$ are tight for all possible r, d, and n by analyzing the following graph: assume for simplicity that $(d+1)|n$ and let $k = n/(d+1)$. For $i = 1, \dots, k$, let H_i be the complete graph (without self-loops) on $(d+1)$ vertices with a single edge (u_i, v_i) removed. Let G be the d-regular simple graph formed by taking a disjoint union of $H_1, \dots, H_k$ and then connect them in a cycle by adding the edges $(v_1, u_2), (v_2, u_3), \dots, (v_{k-1}, u_k), (v_k, u_1)$.

Problem 2.(Interlacing under edge removal) We have seen that when we remove a *vertex* from a graph, the eigenvalues of the graphs' *adjacency matrices* interlace each other. Now we will see that when we remove an *edge* from a graph, the eigenvalues of the graphs' *Laplacians* interlace each other.

Specifically, let G be an undirected, possibly weighted graph, and let G' be obtained by removing a single edge from G (i.e. setting its weight to zero). Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n-1} \leq \lambda_n$ be the eigenvalues of the Laplacian of G, and let $0 = \lambda'_1 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda'_n$ be the eigenvalues of the Laplacian of G'. Prove that:

$$\lambda'_1 \leq \lambda_1 \leq \lambda'_2 \leq \lambda_2 \leq \cdots \leq \lambda'_n \leq \lambda_n.$$

(Hint: if $e = (a, b)$ is the removed edge, and $S \subseteq \mathbb{R}^n$ is a k-dimensional subspace, then $S \cap (\delta_a - \delta_b)^\perp$ has dimension at least $(k-1)$.)

Problem 3.(Directed graphs) Let G be a (possibly weighted) Eulerian digraph with normalized Laplacian $N = I - A$, where $A = D^{-1/2}MD^{-1/2} = D^{-1/2}WD^{1/2}$ is the *normalized adjacency matrix*. Define $\alpha_1, \ldots, \alpha_n$ by

$$\alpha_k = \max_{S \subseteq \mathbb{R}^n : \dim(S) = k} \min_{x \in S - \{0\}} \frac{x^T A x}{x^T x} = 1 - \min_{S \subseteq \mathbb{R}^n : \dim(S) = k} \max_{x \in S - \{0\}} \frac{x^T N x}{x^T x}.$$

and $\beta_1, \ldots, \beta_n$ by

$$\beta_k = \max_{S \subseteq \mathbb{R}^n : \dim(S) = k} \min_{x \in S - \{0\}} \frac{\|Ax\|^2}{\|x\|^2}.$$

In both cases the minimum is taken over subspaces S of $\mathbb{R}^n$.

1. The α's and β's are the eigenvalues of the normalized adjacency matrix of certain *undirected* graphs associated with G. Describe these graphs. (Your descriptions can be either combinatorial or in terms of random walks.)
2. Calculate the α's and β's for the directed cycle on n vertices.
3. We know that when G is undirected, $\alpha_2 < 1$ iff G is connected. Show that the same is true for general Eulerian G.
4. Given an example of an undirected and connected graph G where $\beta_2 = 1$.
5. When A is normal, give formulas for the α's and β's in terms of the (possibly complex) eigenvalues of A.

Problem 4.(Cayley graphs vs. Error-Correcting Codes) An *error-correcting code* is a function $\text{Enc} : \{0,1\}^k \to \{0,1\}^n$ that allows us to encode a k-bit message m into a longer, n-bit codewords $c = \text{Enc}(m)$, such that even if c is corrupted by a bounded amount of noise, it is still possible to decode and recover the correct message m. Specifically, we wish to maximize the *minimum distance* $\Delta = \min_{m \neq m'} d_H(\text{Enc}(m), \text{Enc}(m'))$, where $d_H(c, c')$ denotes the *Hamming distance* between c and c', defined to be the number of coordinates in which c and c' differ. It can be shown that we can always correctly decode a codeword that has been corrupted in at most e positions if and only if $d > 2e$. Typically, the goal when designing error-correcting codes is to design codes for arbitrarily large *blocklength* k that maximize both the relative minimum distance Δ/n as well as the *rate* k/n.

Most explicit error-correcting codes are *linear*, meaning that $\text{Enc}(m) = Am \bmod 2$, for an $n \times k$ $\{0,1\}$ matrix A, where the mod 2 is applied componentwise. Let $S \subseteq \{0,1\}^k$ be the set of rows of A, and let G be the graph $\text{Cay}(\mathbb{Z}_2^k, S)$. Observe that the degree of G is exactly equal to the blocklength n of the code. Prove that the minimum distance of the code is exactly equal to $\lambda_2/2$, where λ_2 is the second-smallest eigenvalue of the Laplacian of G.

Thus, the problem of designing good linear error-correcting codes (over $\mathbb{Z}_2$) is equivalent to the problem of constructing Cayley graphs for the group $\mathbb{Z}_2^k$ whose degree is small relative to k and whose second eigenvalue is large relative to k.